Last one

Thursday, 29 April 2021 10:04 AM

Overview of MMSE, LLSE, Jointly Gaussian RVs

- Hilbert projection theorem — base → applied to estimators
- Orthogonality principle (linear and mmse)
- Jointly Gaussian RV lemmas

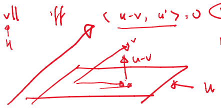
$\mathbb{R}^n$ w/ inner product

H , $U \subseteq H$

$\forall v \in H$, $\exists$ a unique $\omega \in U$ that's closest in U to v

$\omega = \arg\min_{u \in U} \| u - v \|$ iff $\langle u - v, u' \rangle = 0$ $\forall u' \in U$

$H = \mathbb{R}^3$



$L[X|Y] = \arg\min_{u \in U = \text{linear fns of } Y} \| u - X \|$ ← $L[X|Y]$ ← point in U (subspace) closest to $X \in H$

$U = \mathcal{L}(Y)$

iff $\langle X - L[X|Y], u' \rangle = 0$, $u' \in U = \mathcal{L}(Y)$ ←

⟺ $E[(X - L[X|Y]) u'] = 0$ $\forall u' \in U = \mathcal{L}(Y)$

$u' = a + bY$, some a, b

$\forall a, b$ - $E[(X - L[X|Y])(a + bY)] = 0$ ← ①

⇒ ① LLSE 'unbiased'

② Error of estimation ⟂ orthogonal to observations

$E[L[X|Y]] = E(X)$

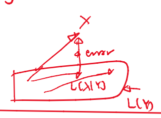
① set $a = 1, b = 0$ in ①

$E[(X - L(X|Y))(1)] = 0$

⟺ $E(X) = E(L(X|Y))$ ←

② set $a = 0$, $b = 1$

$E[\underbrace{(X - L(X|Y))}_{\text{error}} Y] = 0$



RVs X, Y are JG

if $\forall a, b \in \mathbb{R}$

$aX + bY \sim$ normal RVs

① X, Y are JG & uncorrelated

⇒

X, Y indep

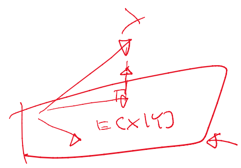
② if X, Y are JG, then

$MMSE(X|Y) = L[X|Y]$

$= E(X|Y)$

$\int x f_{X|Y}\, dx$

proof in HW



$E(X|Y)$ ⇒ define $U = G(Y)$ subspace of all fncs of Y

closest pt on $G(Y)$ to X — MMSE

if $E[(X - E[X|Y]) g(Y)] = 0$, $\forall g$ in functions

MMSE for Jointly Gaussian Random Variables

Provide justification for each of the following steps (1 - 5) to prove that the LLSE is equal to the MMSE estimator for jointly Gaussian random variables X and Y.

Let $g(X) = L[Y \mid X]$.

$$
\begin{aligned}
& E[(Y - g(X))X] = 0 && (1)\\
\implies\ & \operatorname{cov}(Y - g(X), X) = 0 && (2)\\
\implies\ & Y - g(X) \text{ is independent of } X && (3)\\
\implies\ & E[(Y - g(X))f(X)] = 0\ \forall f(\cdot) && (4)\\
\implies\ & g(X) = E[Y \mid X] && (5)
\end{aligned}
$$

$L(Y|X) =$

1. Gaussian Estimation

Let $Y = X + Z$ and $U = X - Z$, where X and Z are i.i.d. $\mathcal{N}(0, 1)$.

(a) Find the joint distribution of U and Y.

(b) Find the MMSE of X given the observation Y, call this $\hat{X}(Y)$.

(c) Let the estimation error $E = X - \hat{X}(Y)$. Find the conditional distribution of E given Y

a) $f(U = u, Y = y) = f_U(u) f_Y(y) = \ldots$

$f_{U,Y}(u, y) = f_{X,Z}\left(\frac{u+y}{2}, \frac{y-u}{2}\right) = \ldots$

U, Y are JG & uncorr ⇒ indep.

$aU + bY \sim N(\cdot, \cdot)$ ⇒ JG

$Cov(U, Y) = E(UY) - E(U)E(Y)$

$= E(X^2 - Z^2)$

$= E(X^2) - E(Z^2) = 0$

$U \sim X - Z \sim N(0,$

$Var(U) = Var(X) + Var((-1) \times Z)$

$= 1 + (-1)^2 Var(Z)$

$= 1 + 1 = 2$

b) X, Y are JG ⇒ $M(X|Y) = L(X|Y) = E(X) + \frac{Cov(X, Y)}{Var(Y)}(Y - E(Y))$

∵ JG

$= 0 + \frac{Cov(X, X + Z)}{Var(Y)}(Y - 0)$

"bilinearity"

$= \frac{Cov(X, X) + Cov(X, Z)}{Var(Y)}$;

$= \frac{Var(X)}{Var(Y)} Y = \frac{1}{2} Y$

$U = X - Z \sim N(0, 2)$

c) $f_{E|Y}(e|y) = \frac{f_{E,Y}}{f_Y} =$ ______

$E = X - E(X|Y) = X - \left(\frac{1}{2}Y\right) = X - \frac{1}{2}(X + Z) = \frac{1}{2}X - \frac{1}{2}Z$

$= \frac{1}{2}U \sim$

⇒ $\frac{1}{2}U \sim N\left(0, \frac{1}{2}\right)$

∴ U & Y are indep (from a) working)

$E \sim N\left(0, \frac{1}{2}\right)$

$f_{E|Y}(e|y) = f_E(e) =$

Goal of K-filter

$\text{MMSE}(X_n \mid Y_1, Y_2, \ldots, Y_n)$

state space equations

$X_n = AX_{n-1} + Z$ ← gaussian noise

$Y_n = BX_n + W$ ← gaussian noise

2. Overlapping Normals $L(X_n \mid Y_1, \ldots, Y_n)$

As you will see in the lab, a big part of the Kalman filter is "overlapping" two normal distributions. In particular, suppose at the current time step, you have a state $X \sim \mathcal{N}(\mu, \sigma_1^2)$ and an observation $Y \sim X + \mathcal{N}(0, \sigma_2^2)$. The two noises are independent.

$X = \mu + N(0, \sigma_1^2)$

(a) Not knowing Y, what is your best guess of X?

(b) X and Y are jointly Gaussian. Write the vector $[X\ Y]^T$ as an affine transformation of independent unit Gaussians. i.e find $A \in \mathbb{R}^{2\times 2}$ and $\mu \in \mathbb{R}^2$ such that $[X\ Y]^T = AZ + \mu$ where $Z = [Z_1\ Z_2]^T$ and $Z_1, Z_2 \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$.

(c) What is $\mathbb{E}[X \mid Y]$?

(d) Given $Y = y$, the conditional distribution $X \mid Y = y$ turns out to be normal, which is the reason why we can continue using the same process for future time steps. In fact, it turns out $\text{var}(X \mid Y)$ is constant, so it's always equal to its expectation $\mathbb{E}[\text{var}(X \mid Y)] = \mathbb{E}[(X - \mathbb{E}[X \mid Y])^2]$. WLOG assume $\mu = 0$ and show that this is equal to $\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}$.

assumption: noise not correl. X; Y

← def^n of conditional variance

a) $E(X) = \mu$

b) $\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} \sigma_1 & 0 \\ \sigma_1 & \sigma_2 \end{bmatrix} \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} + \begin{pmatrix} \mu \\ \mu \end{pmatrix}$

$X \sim Y - N(0, \sigma_2^2)$ ← re-arrange

$X \mid Y_y \sim y - N(0, \sigma_2^2)$

$\text{Var}(X) = \text{Var}(\sigma_1 Z_1) = \sigma_1^2 I = \sigma_1^2$

$\text{Cov}(X, X + N(0, \sigma_2^2)) = \text{Var}(X) + 0 = \sigma_1^2$

c) X, Y are JG ⇒ mmse = llse

$E(X \mid Y) = L(X \mid Y) = E(X) + \frac{\text{Cov}(X,Y)}{\text{Var}(Y)}(Y - E(Y))$

$= \mu + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}(Y - \mu)$

d) $E((X - L(X \mid Y))^2)$

$= E\left(\left(X - \frac{\sigma_1^2}{\sigma_1^2+\sigma_2^2} Y\right)^2\right)$

$Y \sim X + N(0, \sigma_2^2) = W$

$Y \sim X + W$

$W \sim N(0, \sigma_2^2)$ ← noises are ind from Q

$X \sim \mu + N(0, \sigma_1^2)$

$= E\left[\left(X - \left(\frac{\sigma_1^2}{\sigma_1^2+\sigma_2^2}\right)(X + W)\right)^2\right]$

$= E\left(\left(\frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2} X - \frac{\sigma_1^2}{\sigma_1^2+\sigma_2^2} W\right)^2\right)$

$= E\left(\left(\frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2} X\right)^2\right) + E\left(\left(\frac{\sigma_1^2}{\sigma_1^2+\sigma_2^2}\right)^2 W^2\right) - 2(\ )(\ )\, E(XW)$

$E(XW) = E(X)\,E(W) = 0$

$= \left(\frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}\right)^2 E(X^2) + \left(\frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}\right)^2 E(W^2)$

∵ $\mu = 0$, $\text{Var}(X) = E(X^2)$

$E(X) = \mu = 0$

∵ $E(W) = 0$

$E(W^2) = \text{Var}(W)$

$X \sim N(\mu = 0, \sigma_1^2)$, $W \sim N(0, \sigma_2^2)$

$= \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2}$

$= 0$